



# MISO LIMS

HTS Management by Earlham Institute and  
Ontario Institute for Cancer Research (OICR)

# MISO Purpose

- ▶ Track samples processing into libraries
  - ▶ Follow samples (DNA extracts) through library preparation and pooling for sequencing
- ▶ Collect metadata for SRA submission
- ▶ Monitor sequencer output and provide notifications and metrics
  - ▶ Allow use of different sequencing platforms, including Illumina, PacBio, 454, and Solid
- ▶ Enable inter-team communication and high-level reporting
- ▶ Provide a mature, open and extendable alternative for lab tracking

# Background

MISO is a inventory tracking system for HTS sequencing labs.  
Currently developed in a highly coordinated cross-institute effort.

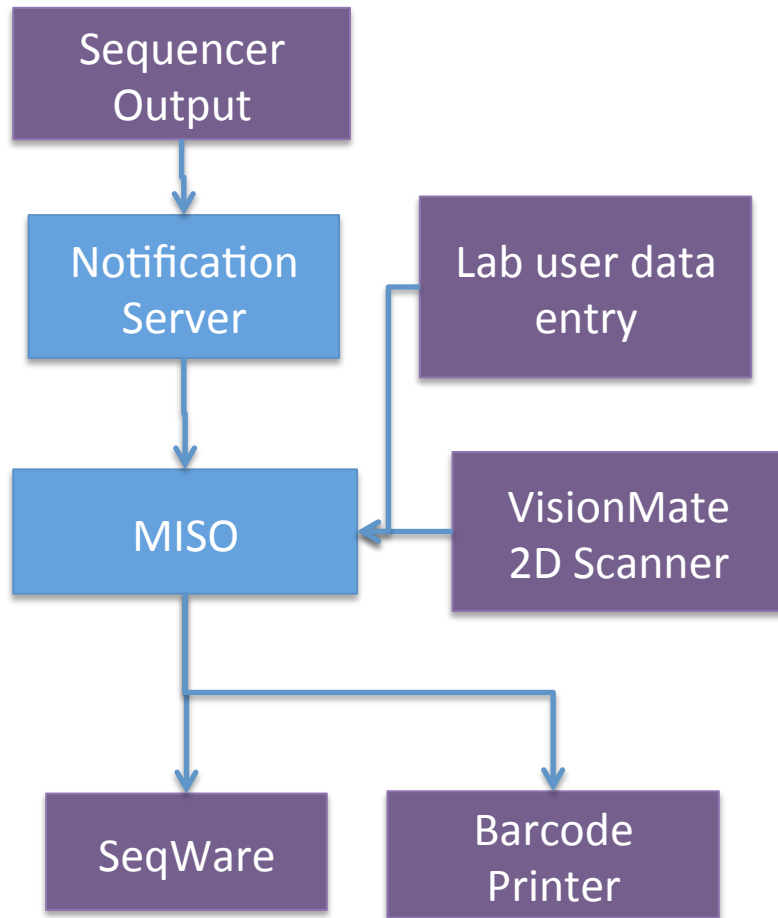
OICR has a HTS sequencing lab

- ▶ Managed through Geospiza GeneSifter LE LIMS
- ▶ Geospiza was getting too expensive and feature requests weren't being added
- ▶ Geospiza is proprietary software and thereby inhibiting to improvements to workflows for lab users

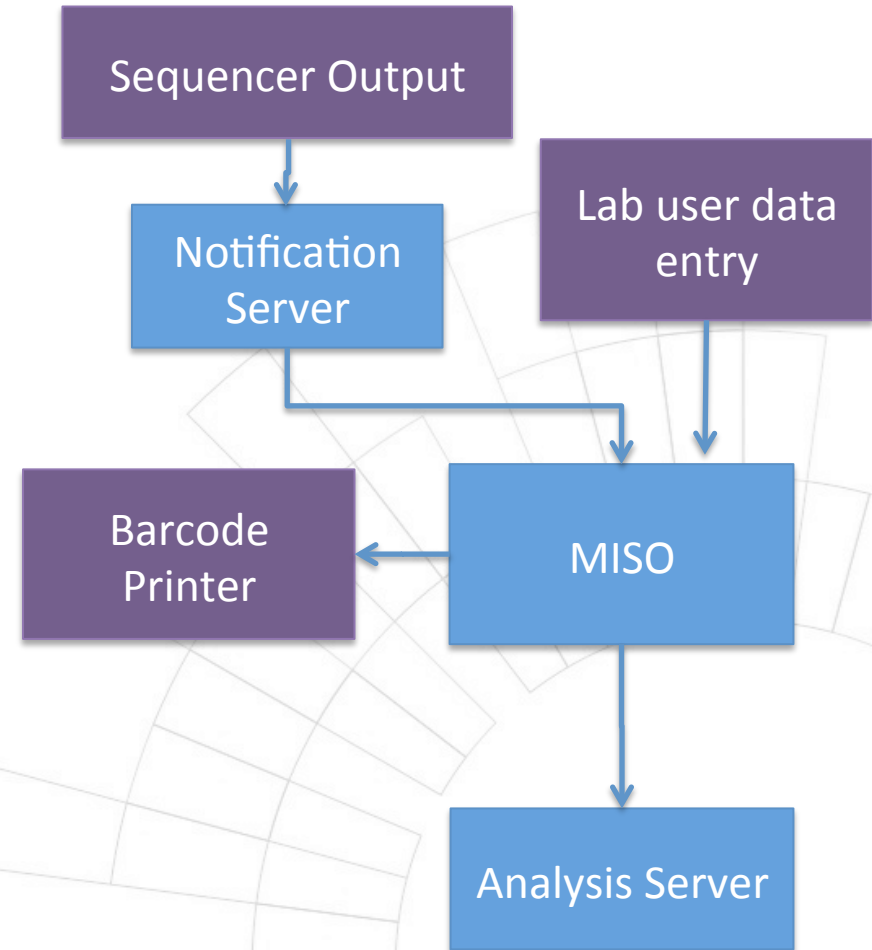
Earlham (formerly TGAC) has a HTS sequencing lab

- ▶ Developed in house LIMS (previously presented)
- ▶ Mostly small feature and maintenance work
- ▶ OICR manager visited and...collaboration!

# Architecture



OICR



Earlham



Together? Forever?

# But It Works on My Machine

- ▶ OICR and Earlham have similar, but distinct needs
  - ▶ Earlham mostly handles DNA extracts, OICR handles tissue samples, slides, and extracts
  - ▶ Similar set of instruments, but OICR mainly uses LIMS to track Illumina, Earlham Illumina and PacBio
  - ▶ OICR has legacy data in Geospiza
- ▶ OICR did not want to fork MISO entirely
  - ▶ We worked out a development plan to integrate OICR features into Earlham's mainline
  - ▶ Some are useful, some are irrelevant to Earlham

# Social Aspects

MISO now actively developed by Earlham and OICR

- ▶ 1 developer at Earlham
- ▶ 4 developers at OICR (until Geospiza sunset)
- ▶ All active work visible on GitHub
- ▶ Code reviews for everything with at least two developers
- ▶ All features integrated into mainline (site-specific repositories are mostly configuration)

A faint, stylized background illustration of a hand holding a fan of cards. The hand is positioned at the top right, with fingers spread, holding several cards that fan out towards the left. The cards are represented by simple rectangular outlines with some internal lines suggesting a grid or text layout. The entire graphic is rendered in a light gray color, creating a subtle watermark effect.

# New Developments



# New Features

- ▶ Tissue processing workflow
  - ▶ Bulk input via Handsontable (Excel-like web interface)

Save														
	Library Alias	Sample Alias	Description	Matrix Barcode	Platform	Type	Selection	Strategy	Barcode Kit	Index 1	Index 2	Kit	Volume	Conc.
1		MGP_0003_Bn_P_nn_1-1_D_S1			Illumina	Mate Pair	RT-PCR	AMPLICON	TruSeq Single Index	Index 01 - ATCACG		TruSeq RNA Access	0.0	0.00
2		MGP_0002_Bn_R_nn_1-1_D_S1			PacBio	2kb Shotgun	Hybrid Selection	WGS	No barcode			KAPA Hyper Prep	0.0	0.00
3		MGP_0001_Bn_R_nn_1-1_D_S1			Illumina	Single End	DNase	CLONE	Nextera Dual Index	N701 - TAAGGCGA	N501 - TAGATCGC	Nextera DNA	0.0	0.00

# New Features

- ▶ Storage boxes for handling frozen barcoded tubes
- ▶ Scanner integration in the UI

Box Use: DNA

Box Size: 8 × 12 (can be scanned by your lab's bulk scanner)

**Contents** Options ▾

	1	2	3	4	5	6	7	8	9	10	11	12
A	○	●	●	●	○	○	○	○	○	○	○	○
B	○	○	○	○	○	○	○	○	○	○	○	○
C	○	○	○	○	○	○	○	○	○	○	○	○
D	○	○	○	○	○	○	○	○	○	○	○	○
E	○	○	○	○	○	○	○	○	○	○	○	○
F	○	○	○	○	○	○	○	○	○	○	○	○
G	○	○	○	○	○	○	○	○	○	○	○	○
H	○	○	○	○	○	○	○	○	○	○	○	○

B10: Empty

Selected Position:

Name:

Alias:

Barcode:  Lookup

Update Position

Remove Tube Trash Tube

Show 96 entries

Position ▲	Element Name ◇	Element Alias ◇
A02	<a href="#">SAM32</a>	<a href="#">MGP_0001_Bn_R_nn_1-1_D_1</a>
A03	<a href="#">SAM33</a>	<a href="#">MGP_0003_Bn_R_nn_1-1_D_1</a>
A04	<a href="#">SAM14</a>	<a href="#">MGP_0001_Bn_P_nn_1-1_D_S1</a>

Showing 1 to 3 of 3 entries

Search:  List all Box Contents

# New Features

- Order management
  - Track what needs to be sequenced with what chemistry

Orders

Partitions:

Platform:

Sequencing Parameters:

Save

Partitions:

Platform:

Sequencing Parameters:

Add

Order Completion

Platform	Parameters	Failed	Requested	Remaining
Illumina - Illumina HiSeq 2500	v4 2x126	2	0	0
Illumina - Illumina HiSeq 2500	v3 2x101	0	1	1

Unfulfilled Pool Orders						
Show <input type="text" value="25"/> entries		Search: <input type="text"/>				
Pool Name	Alias	Platform	Sequencing Parameters	Requested	Remaining	
<a href="#">IPO5</a>	<a href="#">TEST pool 4</a>	Illumina - Illumina HiSeq 2500	Rapid Run 2x51	2	2	
<a href="#">IPO3</a>	<a href="#">TEST pool 1</a>	Illumina - Illumina HiSeq 2500	v3 2x101	1	1	
<a href="#">IPO1</a>	<a href="#">asdfasd</a>	Illumina - Illumina HiSeq 2500	v4 2x126	3	3	
Showing 1 to 3 of 3 entries						

# New Features

## ► Front-end input validation

Oh snap!  
This form seems to be invalid

Library Information

Library ID: 8

Location:

Name: LIB8

Parent Sample: [MGP\\_0001\\_Bn\\_R\\_nn\\_1-1\\_D\\_1 \(SAM32\)](#)

Alias:\* MGP\_0001\_bn\_R\_PE\_301\_WG

Must match (A-Z)(3-5) (0-9)(3-4)(0-9)(CR)(0-9)(1-2) (nn)(A-Z)(1)(a-z)(1) ((nRPXMCFETO)) (SE|PE|MP|?) (7) (nn)(d(2,6))dk(1-15)(EX|CH|BS|WG|TR|WT|SM|MR)(X?)

Description:\* test

77  
251

## ► Change logs

Changes	
Summary	Time
admin has changed: ready: 1 → 0	2016-06-17 22:08:08.0
admin has changed: QC passed: n/a → 1, ready: 0 → 1	2016-06-07 14:09:07.0
admin1 has changed: ready: 1 → 0	2016-04-25 19:24:38.0
admin created pool.	2016-04-25 19:19:40.0
admin Added: LDI1	2016-04-25 19:19:40.0

# New Procedures

Shared development

- ▶ *Everything* gets code reviewed
- ▶ Improved testing and plans for more

More plug-and-play

- ▶ Externalization of configuration
- ▶ Weekly releases at OICR
- ▶ Fewer Earlham-specific assumptions

# Future Plans

## Tissue processing workflows

- ▶ Currently, OICR-specific; Earlham intends to use, so will be generalised
- ▶ OICR has variable workflow, so this will expand anyway

## Better UI

- ▶ Excel-like Handsontable used in new developments, will be applied to existing data import
- ▶ Customisable reporting

# Availability

- ▶ Source available on GitHub
- ▶ Docker container available for low-effort testing
- ▶ On my laptop! Find me and request a demo

<https://github.com/TGAC/miso-lims>

<http://miso-lims.slack.com>

New institutes welcome.  
We're developing a community.

# Team



Rob Davey (Earlham)



Tony DeBat (OICR)



Chris Salt (Earlham)



Andre Masella (OICR)



Heather Armstrong (OICR)



Dillan Cooke (OICR)



Morgan Taschuk (OICR)





<http://earlham.ac.uk/davey-group>



<http://bit.do/oicrgsi>

Funding provided by:

